DOCUMENT RESUME

ED 308 243 TM 013 663

AUTHOR Eggen, Theo J. H. M.; van der Linden, Wim J.
TITLE The Use of Models for Paired Comparisons with Ties. Project Psychometric Aspects of Item Banking No. 8. Research Report 86-8.
INSTITUTION Twente Univ., Enschede (Netherlands). Dept. of Education.
PUB DATE 86
NOTE 30p.
AVAILABLE FROM Mediatheek, Department of Education, University of Twente, P.O. Box 217, 7500 AE Enschede, The Netherlands.
PUB TYPE Reports - Evaluative/Feasibility (142)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS *Comparative Analysis; Equations (Mathematics); Estimation (Mathematics); Mathematical Models; *Models; *Research Methodology; Statistical Analysis
IDENTIFIERS *Paired Comparisons; Parameter Identification; Rasch Model; *Tie Parameters

ABSTRACT

In experiments with paired comparisons, judges are occasionally allowed to express indifference between alternatives. For the analysis of such data, models for paired comparisons with ties are needed. Several models with tie parameters are reviewed. All are extensions of the basic models of L. L. Thurstone (1927) and R. A. Bradley and M. E. Terry (1952). Some of these models can be reparameterized into an exponential family of distributions. This gives the attractive possibility of conditioning individual tie parameters out of the model so that the object parameters can be estimated independently of the tie parameters and with desirable asymptotic properties. In addition, focus is on the use of models with tie parameters for analyzing induced experiments with ties in which the original experiment has no ties, or even no comparisons. This technique may have considerable advantages over analyzing the original experiment, as is shown in the case of the Rasch model for binary test items. One data table and one figure are provided. (Author/SLD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED308243

The Use of Models for Paired Comparisons with Ties

Research Report 86-8

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

J. NELISSEN

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

Theo J.H.M. Eggen
Wim J. van der Linden

TM013643

Department of Education

Division of Educational Measurement and Data Analysis

University of Twente

Project Psychometric Aspects of Item Banking No. 8

Colophon

Typing : A. Burchartz, L. Padberg
Cover design : M. Driessen, AV-section, University of Twente
Printed by : Central Reproduction Department, University of Twente

The Use of Models for Paired Comparisons with Ties

Theo J.H.M. Eggen

Wim J. van der Linden

To appear in W.E. Saris & I.N. Gallhofer (Eds.), Sociometric research: Volume 1. Data collection and scaling. London: MacMillan.

Abstract

In experiments with paired comparisons judges are occasionally allowed to express indifference between alternatives. For the analysis of such data, models for paired comparisons with ties are needed. In this paper several models with tie parameters are reviewed, which are all extensions of Thurstone's (1927) and Bradley-Terry's (1952) basic models. Some of these models can be reparameterized into an exponential family of distributions. This gives the attractive possibility of conditioning individual tie parameters out of the model so that the object parameters can be estimated independently of the tie parameters and with desirable asymptotic properties. In addition, the paper pays special attention to the use of models with tie parameters for analyzing induced experiments with ties in which the original experiment has no ties, or even no comparisons. This may have considerable advantages over analyzing the original experiment, as is shown for the case of the Rasch model for binary test items (Rasch, 1960).

Introduction

The method of paired comparisons has a long and wealthy tradition which has been caused not only by its simplicity and statistical elegance but also by its wide applicability. Among the most important applications are its use in such areas as taste testing, personnel rating, consumer testing, the psychological study of preference and choice behavior, attitude measurement, and tournaments.

Although the method was used first by Fechner (1860) in psychophysics, it can be stated that the introduction of probability models for comparisons of objects by Thurstone (1927) and Zermelo (1929), and in particular the independent discovery of Zermelo's model by Bradley and Terry (1952), were the starting points of an extensive series of publications on the method of paired comparisons. Valuable reviews of the literature are given in David (1963) and Bradley (1976), while Davidson and Farquhar (1976) provide an extensive bibliography.

In the original papers, the attention was focused mainly on designs with an equal number of repetitions for each comparison and for which independence of all comparisons could be assumed. For this case, procedures for parameter estimation and test of model fit were given by Mosteller (1951) and Bradley and Terry (1952). Nowadays the situation has changed completely. Not only have alternative models been developed (e.g., by Thomson & Singh, 1967), but also a great many extensions of the basic models and designs have

become available. Some of these are models for unbalanced designs - i.e. unequal numbers of repetitions of comparisons - (Dijkstra, 1960), models that account for effects of the order of presentation of the objects in a pair (Beaver & Gokhale, 1975; Davidson & Beaver, 1977; Kousgaard, 1979), and models permitting the occurence of ties in comparisons (Glenn & David, 1960; Rao & Kupper, 1967; Davidson, 1970; Kousgaard, 1976). Furthermore, the theory of paired comparisons has been generalized to triple comparisons, factorials in comparisons, and multivariate paired comparisons (see, e.g., Bradley, 1976, and the references therein). Some statistical tests for dependence between comparisons have been provided by Molenaar (1973). Finally, graph-theoretic and combinatorial studies on tournaments have provided all kinds of designs for paired comparison experiments.

In this paper, we restrict ourselves to models for paired comparisons with ties. In the first part of the paper a review of models with special parameters for ties will be given. In particular attention will be payed to extensions of the Bradley-Terry model and to models with individual tie parameters. This part of the paper also deals with statistical issues in parameter estimation and shows how the method of conditional maximum-likelihood can be used to remove the individual tie parameters from the estimation equations so that consistent estimates of the object parameters are obtained. The second part of the paper deals with a special application of models with tie parameters. As already noted by Gridgeman (1959), an obvious advantage of such models is that they can handle

more response formats and are able to distinguish between abilities of respondents or judges. It is the contribution of this paper to point at a less obvious application of models with tie parameters, namely the analysis of completely different experiments and models for which it can be shown that a transformation of the data exists that induces a paired-comparisons experiment with ties. The original experiment does not need to have ties or even comparisons. We will illustrate this with the analysis of binary test items with the Rasch (1960) model. Analysis of the induced experiment instead of the original one may have several advantages. An example with the Rasch (1960) model for binary test items will illustrate this.

Models with Tie Parameters

Thurstone's (1927) basic model applies to the following situation. There are n objects (treatments, stimuli, etc.) presented in pairs to respondents who are asked to express preference for one object of the pair with respect to some attribute. The model assumes that the probability of preferring object i to j is governed by

$$(1) \qquad P(X_i > X_j) = (2\pi)^{-1/2} \int_{-(u_i-u_j)}^{\infty} \exp(-y^2/2)\, dy = \Phi(u_i-u_j), \qquad i \neq j = 1,\ldots,n$$

where X_i is the random variable denoting the response to object i ($i = 1,\ldots,n$) with location u_i on the attribute and $\Phi(.)$ denotes

the standard normal distribution function.

Replacing the normal density in (1) by the logistic, the Bradley-Terry model is obtained:

$$(2) \qquad P(X_i > X_j) = \frac{1}{4} \int_{-(\ln\pi_i - \ln\pi_j)}^{\infty} \operatorname{secn}^2(y/2)\, dy$$

$$= \Psi(\ln\pi_i - \ln\pi_j)$$

$$= \pi_i/(\pi_i+\pi_j), \qquad i \neq j = 1,\ldots,n$$

now with π_i as a parameter representing object i's location and with $\pi_i > 0$, $\sum_{i=1}^{n} \pi_i = 1$, and where $\Psi(z) = (1+e^{-z})^{-1}$ is the logistic distribution function.

Glenn and David (1960) were the first to present a modification of the Thurstone model that makes provision for the possibility for judges to state an indifference between objects. These indifferences are called ties. In Thurstone's model in (1) it is assumed that differences between the responses X_i and X_j, however small, are discernible by the respondent who, consequently, is always able to state a preference. Glenn and David, however, postulated that when the absolute difference lies below a certain threshold value, i.e. when $X_i - X_j$ lies between $-\tau$ and τ, a respondent will not be able to give a preference and a tie occurs. For normally distributed response variables X_i with expectation u_i and variance 1, Glenn and David's model follows as in Figure 1.

Insert Figure 1 here

Algebraically, the model can be given as:

$$P(i|i,j) = P(X_i - X_j > \tau) = \Phi(u_i - u_j - \tau)$$

$$(3) \qquad P(j|i,j) = P(X_i - X_j < -\tau) = \Phi(u_j - u_i - \tau)$$

$$P(0|i,j) = P(-\tau < X_i - X_j < \tau) = \Phi(u_i - u_j + \tau) - \Phi(u_i - u_j - \tau),$$

where, respectively $P(.\ i,j)$ are the probabilities that i is preferred, j is preferred, or a tie occurs, respectively. It should be noted that τ, the treshold or tie parameter, can be interpreted as the judges' ability to discriminate between objects, and that Thurstone's model is obtained for $\tau = 0$. In fact, Glenn and David did not propose the model in (3) but a version of it in which $\Phi(.)$ is replaced by an angular distribution with distribution function $f(z) = (1+\sin z)/2$. For their considerations to do so, we refer to their paper (Glenn & David, 1960).

The second model for paired comparisons with ties is the one by Rao and Kupper (1967) obtained from the Bradley-Terry model along the same lines as above:

$$P(i|i,j) = \Psi(\ln\pi_i - \ln\pi_j - \ln\xi) = \pi_i/(\pi_i+\xi\pi_j)$$

(4) $$P(j|i,j) = \Psi(\ln\pi_j - \ln\pi_i - \ln\xi) = \pi_j/(\pi_j+\xi\pi_i)$$

$$P(0|i,j) = \Psi(\ln\pi_i - \ln\pi_j + \ln\xi) - \Psi(\ln\pi_i - \ln\pi_j - \ln\xi)$$

$$= \pi_i\pi_j(\xi^2-1)/(\pi_i+\xi\pi_j)(\pi_j+\xi\pi_i),$$

with ξ denoting the threshold or tie parameter.

The original Bradley-Terry model (2) has the property that the odds ratio is equal to $P(i|i,j)/P(j|i,j) = \pi_i/\pi_j$, by which it satisfies Luce's (1959) choice axiom. Davidson (1970) gives an extension of the Bradley-Terry model retaining this property but also allowing for ties:

$$P(i|i,j) = \pi_i/(\pi_i+\pi_j+\nu(\pi_i\pi_j)^{1/2})$$

(5) $$P(j|i,j) = \pi_j/(\pi_i+\pi_j+\nu(\pi_i\pi_j)^{1/2})$$

$$P(0|i,j) = \nu(\pi_i\pi_j)^{1/2}/(\pi_i+\pi_j+\nu(\pi_i\pi_j)^{1/2}),$$

where ν is the tie parameter. In the model the probability of a tie is proportional with constant ν to $(\pi_i\pi_j)^{1/2}$, which is the mean $(\ln\pi_i+\ln\pi_j)/2$ of the object parameters on a logarithmic scale.

The models discussed so far have the common property that one tie parameter is introduced for all judges. This is a drawback if

judges differ in ability to discriminate between objects, which they obviously will do for some sets of objects. In such cases there is a wish for an individual tie parameter. Kousgaard (1976) presents two models with individual tie parameters, one of which will be ignored here because it contains no separate object parameters. The other one can be shown to be identical to Davidson's model in (5) with ν replaced by a tie parameter ν_k whose value is dependent on respondent k. As will be clear form the remaining part of the paper, this model has many attractive properties.

Comparison of the Models

All models above are generalizations of either the Thurstone or the Bradley-Terry model since one of the two follows upon specifying an appropriate value of the tie parameter (viz., $\tau = 0$, $\xi = 1$, $\nu = 0$, and $\nu_k = 0$ for all judges, respectively). Also, all models behave in a fashion that does not go against our intuition. For example, in all models the probability of a preference is decreasing in the tie parameter but increasing in the distance between the object parameters. To the authors of the paper this means that statistical considerations ought to be decisive in the choice of one of these models.

For Glenn and David's model (weighted) least-squares estimation procedures for the parameters are available but not much is known about the properties of the estimators. This is different for the Rao and Kupper and the Davidson models for which maximum-likelihood

procedures have been studied extensively for a wide variety of designs. Of these two, Davidson's model is to be preferred for several reasons. First, numerical simplifications in the estimation procedure are possible due to the fact that the model still satisfies Luce's choice axiom. Second, unlike Davidson's model, for Rao and Kupper's model the uniqueness of the maximum-likelihood estimates has not been established (cf. Davidson & Beaver, 1977). Third, Davidson's model has, as opposed to Rao and Kupper's, the attractive property that for balanced designs the maximum-likelihood estimator of the object parameter is a monotone increasing function of the number of preferences of the object over all other objects. Four, and most importantly, Davidson's model can be reparameterized into an exponential family of probability distributions, which opens the attractive possibility of conditional inference (Andersen, 1980, ch. 3).

In contrast with Davidson's original model, Kousgaard's version allows for individual difference between judges. As was argued before, this is a most favorable provision since for most sets of objects judges obviously differ in their abilities to discriminate between pairs of them. Therefore, our conclusion is that from the existing models Kousgaard's is the one to be preferred, in particular because of its possibility of parameter estimation along the lines of conditional inference. This will be outlined in the following section.

Estimation in the Model with Individual Tie Parameters

Kousgaard's model can be brought on exponential form by the following reparameterization of (5) with individual parameters v_k:

$$\pi^{1/2} = \exp(\theta_i)$$

$$v_k = \exp(\zeta_k)$$

Formally, we then have random variables X_{ijk} representing the outcome of a comparison between object i and j ($i<j = 1,\ldots,n$) by judge k ($k = 1,\ldots,m$) with possible values -1, 0, and 1 for the events "j is preferred", "a tie", and "i is preferred", respectively which are distributed as

$$(6) \qquad p(X_{ijk} = x) = \frac{\exp\{(\theta_i-\theta_j)x + \zeta_k(1+x)(1-x)\}}{\exp(\theta_i-\theta_j) + \exp(\theta_j-\theta_i) + \exp(\zeta_k)},$$

with parameter space $\{(\zeta_1,\ldots,\zeta_m,\theta_1,\ldots,\theta_n) \in R^{m+n};\ \sum_{i=1}^{n} \theta_i = 0\}$ (Kousgaard, 1976). This model, which is easily verified to belong to the exponential family of distributions, is a model with so-called structural parameters $\{\theta_i;\ i = 1,\ldots,n\}$ and incidental parameters $\{\zeta_k;\ k = 1,\ldots,m\}$. Hence, we have no guarantee that the joint maximum-likelihood estimators of the structural parameters are consistent, for if the number of persons tends to infinity, the

same holds for the number of unknown parameters ζ_k (Neyman & Scott, 1948).

If in cases like this sufficient statistics for the incidental parameters are present, conditional maximum-likelihood procedures are suitable. For the distributions in (6) it is obvious that $U_k = \sum_{i<j} (1+X_{ijk})(1-X_{ijk})n_{ijk}$ is sufficient for ζ_k, where n_{ijk} indicates if judge k has compared object i and j ($n_{ijk}=1$) or not ($n_{ijk}=0$). This statistic is simply the number of ties stated by judge k. By a standard argument it follows that the likelihood function associated with the conditional distribution of $\{X_{ijk};\ i<j=1,\ldots,n,\ k=1,\ldots,m\}$ given $(U_1 = u_1, \ldots, U_m = u_m)$ is equal to

$$(7) \qquad L_c = \frac{\exp(\sum_{i=1}^{n} \theta_i s_i)}{\prod_{k=1}^{m} \phi_{(u_k)}(\theta)},$$

with

$$(8) \qquad \phi_{(u_k)}(\theta) = \sum_{(u_k)} \exp\left(\sum_{i=1}^{n} \theta_i s_i\right),$$

where $s_i = \sum_{k=1}^{m} \sum_{j=1}^{n} X_{ijk} n_{ijk}$ is equal to the number of comparisons in which i is preferred minus the number in which another object is preferred to i and $\sum_{(u_k)}$ denotes summation over all possible outcomes with u_k ties.

The attractive thing about (7) is that it does not contain any

tie parameters and thus allows maximum-likelihood estimation of $\{\theta_i\}$ without any contamination by characteristics of the individual judges. As Kousgaard (1976) has demonstrated, the conditional estimators of $\{\theta_i\}$ given by maximization of (7) consequently have desired asymptotic properties.

The foregoing nicely illustrates the logic of conditional inference but still has a practical disadvantage: the computational burden and possible numerical inaccurateness in determining (8). For larger sets of objects the number of terms in (8) and the propagation of errors arising in their iterative numerical evaluation may even become prohibitive.

An alternative to (7) - (8), noted as an aside by Kousgaard, is not to condition $\{X_{ijk}\}$ on $(U_1=u_1,\ldots,U_m=u_m)$ but to condition each individual random variable X_{ijk} on $U_{ijk} = 0$, i.e. on the event of no tie in the comparison between i and j by judge k. In practical terms this means that instead of aggregating our data and then conditioning on the number of ties per judge, we now first remove the ties and only then aggregate our data to estimate the object parameters. Again, it is a standard derivation to show that the resulting conditional likelihood equation yielded by this procedure is of the product-binomial form with the Bradley-Terry model as success parameter:

$$(9) \qquad L_c' = \prod_{i<j} \left(\frac{\exp\theta_i}{\exp\theta_i+\exp\theta_j}\right)^{s_{ij}} \left(\frac{\exp\theta_j}{\exp\theta_i+\exp\theta_j}\right)^{n_{ij}-s_{ij}} ,$$

in which n_{ij} is the number of comparisons between i and j in which no tie has occurred and s_{ij} is the number of comparisons in which i has been preferred to j. This surprising result, which brings us back to the Bradley-Terry model, again shows how conditioning on proper statistics yields estimation equations for the object parameters that are free from the incidental tie parameters. To solve the likelihood equations from (9), a simple algorithm with known properties given by Ford (1957) can be used.

It is this conditioning at the level of the individual events of ties that has brought us to the application in the following section.

The Rasch Model as a Model for Paired Comparisons with Ties

Suppose an individual k with ability α_k responds to a dichotomous test item i with difficulty δ_i. The Rasch model assumes that the probability of a successful response is governed by

$$(10) \qquad P(X_{ki} = 1) = \frac{\alpha_k}{\alpha_k + \delta_i},$$

with $X_{ki} = 1$ denoting a successful response and $X_{ki} = 0$ a failure. The usual estimation procedures for the difficulty parameters in (10) are of the conditional maximum-likelihood type with conditioning on the number of successes per individual to remove the incidental ability parameters from the model (e.g., Fischer, 1974, ch. 14). This approach, however, runs into numerical problems compara-

ble to those for Kousgaard's likelihood equations (7) - (8) mentioned earlier. For this reason standard programs for Rasch analysis do not accept more than 60 items.

The formal resemblance between these two problems as well as between the Rasch model and the Bradley-Terry model suggested the authors to overcome the numerical difficulties for the Rasch model in the same fashion as illustrated above, namely by conditioning on a sufficient statistic for the incidental ability parameter at the level of the individual response variables. An important difference between the present problem and the one by Kousgaard is that for the response variable in (10) reduction of its sample space results in a degenerate statistic; hence, (non-trivial) sufficient statistics for the ability parameter α_K are not available. However, the response process described by (10) can be conceived of as inducing a paired-comparisons experiment at the level of a single individual responding to a pair of items. As this turns out to be an experiment with ties, an analysis of this experiment instead of the original Rasch experiment has the advantage that conditioning as in Kousgaard's model is possible. How this can be done is indicated in the following. Only an outline of the procedure and some results from the first runs of a new computer program are given. For details and more results, we refer to a companion paper by the same authors (van der Linden & Eggen, 1986).

Suppose an individual k responds to two dichotomous items, i and j. The possible outcomes of this experiment can be conceived of as emanating from a paired comparison between the difficulties of the

two items. The outcome (X_{ik}=1, X_{jk}=0) indicates that according to "judge" k item j is more difficult than i, (X_{ik}=0, X_{jk}=1) indicates the reverse, whereas the outcomes with both items right or wrong indicate that k is indifferent with respect to i and j and that a tie has occurred. From (10), the probabilities of these three events are given by:

$$P_k(1,0) = \alpha_k\delta_j/\{(\alpha_k+\delta_i)(\alpha_k+\delta_j)\}$$

(11) $$P_k(0,1) = \alpha_k\delta_i/\{(\alpha_k+\delta_i)(\alpha_k+\delta_j)\}$$

$$P_k(\text{tie}) = (\alpha_k^2+\delta_i\delta_j)/\{(\alpha_k+\delta_i)(\alpha_k+\delta_j)\}$$

It should be noted that among the set of comparisons between all possible pairs from a test of n items some of the outcomes are dependent on others. Van der Linden and Eggen (1986) show that the number of dependencies is a simple function of the individual's number-right score and present several algorithms for removing dependencies from the data.

Conditioning the response variable for individual k defined in (11) on the event of a non-tie yields the result that

(12) $$P[(1,0)|(0,1),(1,0)] = \delta_j/(\delta_i+\delta_j)$$

So a Rasch experiment at the level of an individual responding to a series of n items can be analyzed of as a Bradley-Terry experiment

between pairs of them, provided one is willing to condition on the outcome of a non-tie and to remove dependencies from the data. The parameters in (12) can then be estimated using, e.g., Ford's (1957) iterative algorithm.

To demonstrate the feasibility of the above procedure for estimating larger sets of items, a provisional computer program PAIRCOM was written that made use of the Ford algorithm and removed the dependencies between comparisons of items on the basis of random selection. The program, for which no attempts have been made yet to speed up numerical procedures, was run against PML, one of the standard programs available for the Rasch model (Gustafsson, 1979). Simulated data for 1,000 individuals with ability parameter values sampled from N(2,1) and items with difficulty parameters also sampled from the same distribution were used. Table 1 shows the results in CPU-seconds. The maximum number of items accepted by

---

Insert Table 1 about here

PML is equal to 60. It appears that PAIRCOM is able to cope with larger sets of items and needs fewer CPU-seconds for the estimation of sets for which comparison with PML is possible. More extensive comparisons and studies with larger sets of items are planned. Table 1 also gives the time needed to edit the data (removal of ties and dependencies). This editing was based on an algorithm with

random selection. Attempts to speed up this algorithm and to study possible alternatives are undertaken in van der Linden and Eggen (1986).

Discussion

The methodology used in the above reformulation of the Rasch model can be summarized as follows: The attention was focussed at an individual responding to a pair of items. On the sample space of this compound experiment a statistic with a range of three possible values was defined, one value indicating dominance of item i over j, another of item j over i, and the remaining one a tie. The probability distribution of this statistic was derived from the original model governing the response of the individual to a single item. Conditioning on the event of a non-tie not only enabled us to get rid of the ties but also freed the model from the characteristics of the individual relevant with respect to the items and permitted independent comparisons between the items. This was made possible because the original model, and hence the model for the pair of items, belonged to the exponential family and the partitioning of the sample space for the pair of items into ties and non-ties was sufficient for the person parameter in the model.

Another possible application of this procedure that comes into mind is the problem of unfolding. Coombs' (1976) observation that preferential choice data are contaminated by the presence of ideal points is basically what in this paper was called the problem of

individual or, following Neyman and Scott's terminology, the problem of incidental parameters. Conditioning on sufficient statistics for individual parameters is therefore a possible stochastic alternative to the deterministic technique of unfolding. The experiment is then reduced to one that yields data belonging to Coombs' quadrant of stimulus-comparison data and is described adequately by the conditional model. A critical step in this procedure is the selection of the model that describes the response of the individual to a single item. The case of the Rasch model dealt with in this paper suggests a choice from the exponential family of distributions.

References

Andersen, E.B. (1980), Discrete statistical models with social science applications. Amsterdam: North Holland Publishing Company.

Beaver, R.J. & Gokhale, D.V. (1975), A model to incorporate within-pair order effects in paired comparisons. Communications in Statistics, 4, 923-939.

Bradley, R.A. & Terry, M.E. (1952), Rank analysis of incomplete block design: I. The method of paired comparisons. Biometrika, 39, 324-345.

Bradley, R.A. (1976), Science, Statistics and paired comparisons. Biometrics, 32, 213-232.

Coombs, C.H. (1976), A theory of data. Ann Arbor: Mathesis press.

David, H.A. (1963), The method of paired comparisons. London: Griffin.

Davidson, R.R. (1970), On extending the Bradley-Terry model to accomodate ties in paired comparison experiments. Journal of the American Statistical Association, 65, 317-328.

Davidson, R.R. & Beaver, R.J. (1977), On extending the Bradley-Terry model to incorporate within-pair order effects. Biometrics, 33, 693-702.

Davidson, R.R. & Farquhar, P.H. (1976), A bibliography on the method of paired comparisons. Biometrics, 32, 241-252.

Dijkstra, O. (1960), Rank analysis of incomplete block designs: a method of paired comparisons employing unequal repetitions on pairs. Biometrics, 16, 176-188.

Fechner, G.T. (1860), Elemente der Psychophysik. Leipzig: Breitkopf und Härtel.

Fischer, G.H. (1974), Einführung in die Theorie psychologischer Tests. Bern, Stuttgart, Wien: Verlag Hans Huber.

Ford, L.R.J. (1957), Solution of a ranking problem from binary comparisons. American Matnematical Monthly, 64, 241-252.

Glenn, W.A. & David, H.A. (1960), Ties in paired comparisons experiments using a modified Thurstone-Mosteller method. Biometrics, 16, 86-109.

Gridgeman, N.T. (1959), Pair comparison, with and without ties. Biometrics, 15, 382-388.

Gustafsson, J.E. (1979), PML: A computer program for conditional estimation and testing in the Rasch model for dicnotomous items. Reports from the Institute of Education University of Göteborg, No. 85.

Kousgaard, N. (1976), Models for paired comparisons with ties. Scandinavian Journal of Statistics, 3, 1-14.

Kousgaard, N. (1979), A conditional approach to the analysis of data from paired comparisons experiments incorporating witnin-pair order effects. Scandinavian Journal of Statistics, 6, 154-160.

Luce, R.D. (1959), Individual choice behavior. New York: John Wiley & Sons.

Molenaar, W. (1973), Paired comparisons: some models and tests. Bulletin of the International Statistical Institute, 15, 212-218.

Mosteller, F. (1951), Remarks on the method of paired comparisons: I. The least squares solution assuming equal standard deviations and equal correlations. Psychometrika, 16, 3-9.

Neyman, J. & Scott, E.L. (1948), Consistent estimates based on partially consistent observations. Econometrica, 16, 1-5.

Rao, P.V. & Kupper, L.L. (1967), Ties in paired-comparison experiments: A generalization of the Bradley-Terry model. Journal of the American Statistical Association, 62, 194-204, corrigenda 63, 1550.

Rasch, G. (1960), Probability models for some intelligence and attainment test. Kopenhagen: Danmarks Paedagogiske Institut.

Thomson, W.A. & Singh, J. (1967), The use of limit theorems in paired comparison model building. Psychometrika, 32, 255-264.

Thurstone, L.L. (1927), Psychophysical analysis. American Journal of Psychology, 38, 368-389.

van der Linden, W.J. & Eggen, T.J.H.M. (1986), The Rasch model as a model for paired comparisons with an individual tie parameter. Enschede, The Netherlands: University of Twente, Department of Education.

Zermelo, E. (1929), Die Berechnung der Turnier-Ergebnisse als ein maximum-problem der Warscheinlichkeitsrechnung. Mathematisch Zeitschrift, 29, 436-460.

Acknowledgments

The authors are indebted to Rien Steen for writing a provisional version the computer program PAIRCOM.

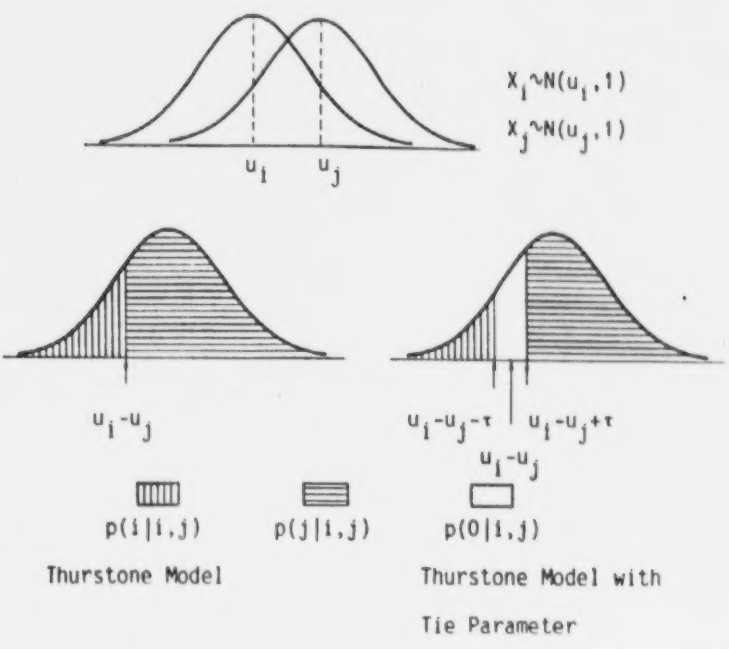

Figure 1. Thurstone model with and without tie parameter in the standard normal density.

Table 1

CPU-Seconds Needed for Comparable Runs with PML and PAIRCOM on Simulated Data for 1,000 Individuals

No. of Items	PML	PAIRCOM	
		Parameter Estimation	Data Editing
20	5.6	1.3	9.8
40	12.6	3.6	13.9
60	31.7	9.8	18.5
100	--	19.7	28.6
200	--	145.0	64.0

Titles of Recent Research Reports

RR-86-1 W.J. van der Linden, The use of test scores for classification decisions with threshold utility

RR-86-2 H. Kelderman, Item bias detection using the loglinear Rasch model: Observed and unobserved subgroups

RR-86-3 E. Boekkooi-Timminga, Simultaneous test construction by zero-one programming

RR-86-4 W.J. van der Linden, & E. Boekkooi-Timminga, A zero-one programming approach to Gulliksen's random matched subtests method

RR-86-5 E. van der Burg, J. de Leeuw, & R. Verdegaal, Homogeneity analysis with k sets of variables: An alternating least squares method with optimal scaling features

RR-86-6 W.J. van der Linden, & T.J.H.M. Eggen, An empirical Bayes approach to item banking

RR-86-7 E. Boekkooi-Timminga, Algorithms for the construction of parallel tests by zero-one programming

RR-86-8 T.J.H.M. Eggen, & W.J. van der Linden, The use of models for paired comparisons with ties

Research Reports can be obtained at costs from Mediatheek, Faculteit Toegepaste Onderwijskunde, Universiteit Twente, P.O. Box 217, 7500 AE Enschede, The Netherlands.

A publication by
the Department of Education
of the University of Twente,
P.O. Box 217,
7500 AE Enschede,
The Netherlands

Department of Education